Full wwPDB X-ray Structure Validation Report ⓘ

Oct 12, 2022 – 07:28 pm BST

PDB ID	:	7QNJ
Title	:	CRYSTAL STRUCTURE OF E.coli ALCOHOL DEHYDROGENASE - FucO MUTANT F254I COMPLEXED WITH FE, NAD+, AND GLYCEROL
Authors	:	Sridhar, S.; Kiema, T.R.; Wierenga, R.K.; Widersten, M.
Deposited on	:	2021-12-20
Resolution	:	1.66 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

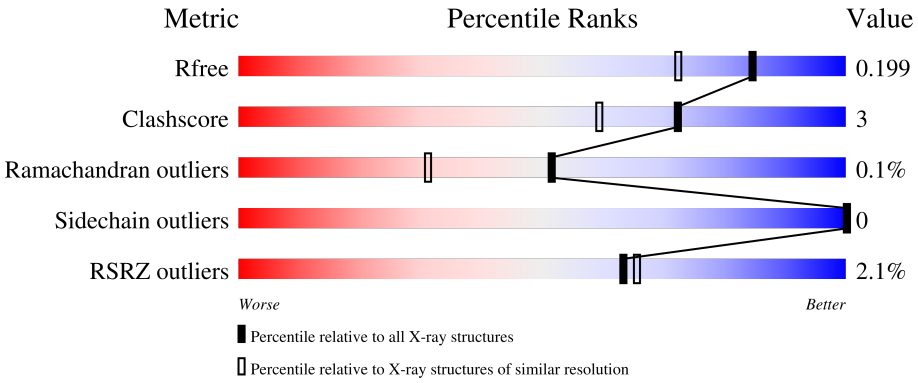
MolProbity	:	4.02b-467
Mogul	:	FAILED
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
buster-report	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.66 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	AAA	390	3% 92% 6%
1	BBB	390	2% 93% 6%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	AAA	405	-	-	X	-
3	GOL	BBB	906	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 12460 atoms, of which 5936 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Lactaldehyde reductase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	383	Total	C	H	N	O	S	88	6	0
			5785	1816	2904	504	545	16			
1	BBB	384	Total	C	H	N	O	S	69	5	0
			5797	1822	2908	504	547	16			

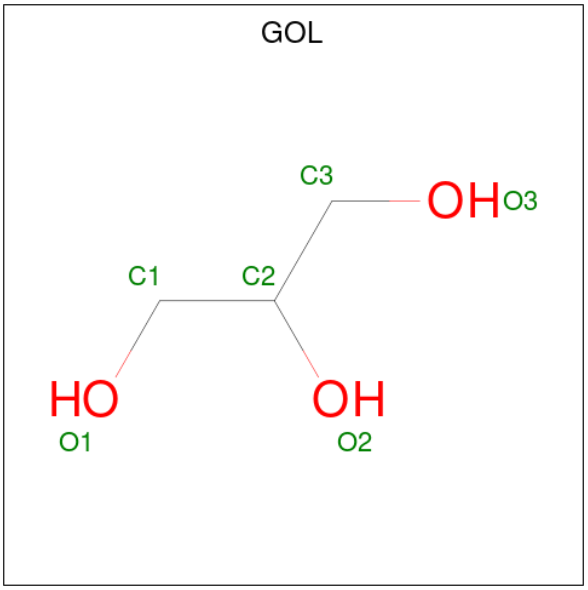
There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	MET	-	initiating methionine	UNP P0A9S2
AAA	254	ILE	PHE	engineered mutation	UNP P0A9S2
AAA	384	THR	-	expression tag	UNP P0A9S2
AAA	385	SER	-	expression tag	UNP P0A9S2
AAA	386	HIS	-	expression tag	UNP P0A9S2
AAA	387	HIS	-	expression tag	UNP P0A9S2
AAA	388	HIS	-	expression tag	UNP P0A9S2
AAA	389	HIS	-	expression tag	UNP P0A9S2
AAA	390	HIS	-	expression tag	UNP P0A9S2
BBB	1	MET	-	initiating methionine	UNP P0A9S2
BBB	254	ILE	PHE	engineered mutation	UNP P0A9S2
BBB	384	THR	-	expression tag	UNP P0A9S2
BBB	385	SER	-	expression tag	UNP P0A9S2
BBB	386	HIS	-	expression tag	UNP P0A9S2
BBB	387	HIS	-	expression tag	UNP P0A9S2
BBB	388	HIS	-	expression tag	UNP P0A9S2
BBB	389	HIS	-	expression tag	UNP P0A9S2
BBB	390	HIS	-	expression tag	UNP P0A9S2

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe) (labeled as "Ligand of Interest" by depositor).

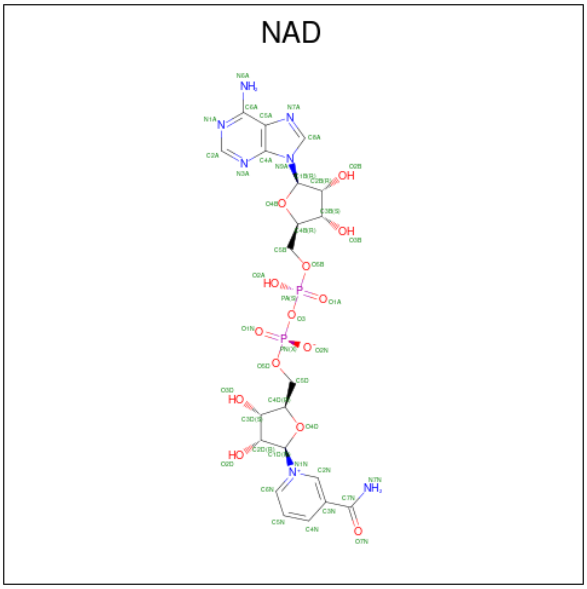
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total Fe 1 1	0	0
2	BBB	1	Total Fe 1 1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total C H O 14 3 8 3	2	0
3	AAA	1	Total C H O 14 3 8 3	2	0
3	AAA	1	Total C H O 14 3 8 3	2	0
3	BBB	1	Total C H O 14 3 8 3	2	0
3	BBB	1	Total C H O 14 3 8 3	2	0
3	BBB	1	Total C H O 14 3 8 3	2	0
3	BBB	1	Total C H O 14 3 8 3	2	0
3	BBB	1	Total C H O 14 3 8 3	2	0
3	BBB	1	Total C H O 14 3 8 3	2	0

- Molecule 4 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: $C_{21}H_{27}N_7O_{14}P_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	AAA	1	Total	C	H	N	O	P	6	0
			70	21	26	7	14	2		
4	BBB	1	Total	C	H	N	O	P	6	0
			70	21	26	7	14	2		

- Molecule 5 is water.

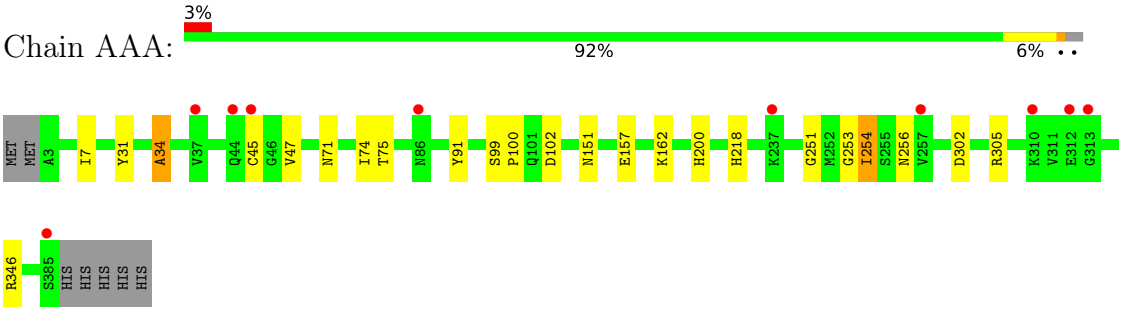
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	280	Total	O	0	0
			280	280		
5	BBB	330	Total	O	0	0
			330	330		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

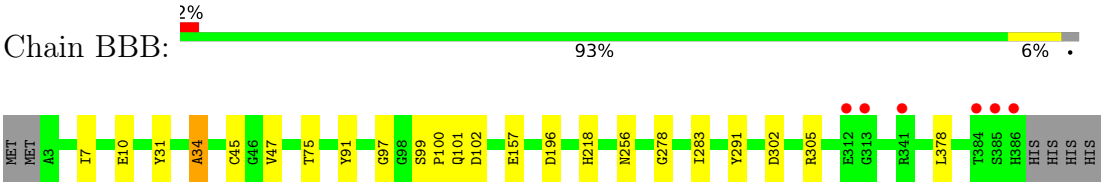
• Molecule 1: Lactaldehyde reductase

Chain AAA:



• Molecule 1: Lactaldehyde reductase

Chain BBB:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	58.14Å 86.22Å 148.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.43 – 1.66 29.41 – 1.66	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.43-1.66) 99.8 (29.41-1.66)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.26 (at 1.66Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.154 , 0.183 0.167 , 0.199	Depositor DCC
R_{free} test set	4492 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	17.5	Xtriage
Anisotropy	0.527	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	12460	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.53% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAD, FE, GOL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	AAA	0.63	0/2952	0.69	0/4016
1	BBB	0.63	0/2952	0.69	0/4017
All	All	0.63	0/5904	0.69	0/8033

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	AAA	0	1
1	BBB	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	34	ALA	Peptide
1	BBB	34	ALA	Peptide

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	AAA	2881	2904	2900	20	0
1	BBB	2889	2908	2898	19	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
3	AAA	18	24	24	4	0
3	BBB	36	48	48	7	0
4	AAA	44	26	26	3	0
4	BBB	44	26	26	1	0
5	AAA	280	0	0	3	0
5	BBB	330	0	0	1	0
All	All	6524	5936	5922	40	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (40) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:7:ILE:H	1:AAA:256:ASN:HD21	1.36	0.71
1:BBB:7:ILE:H	1:BBB:256:ASN:HD21	1.40	0.69
1:BBB:75:THR:H	3:BBB:906:GOL:H32	1.60	0.65
1:BBB:102:ASP:OD2	4:BBB:907:NAD:H2N	1.98	0.63
1:AAA:7:ILE:H	1:AAA:256:ASN:ND2	1.99	0.60
1:BBB:10:GLU:OE2	3:BBB:901:GOL:H12	2.04	0.57
1:AAA:200:HIS:CG	1:AAA:254:ILE:HD11	2.39	0.56
1:BBB:157:GLU:OE1	3:BBB:906:GOL:O3	2.24	0.55
1:AAA:74:ILE:HB	3:AAA:405:GOL:H2	1.88	0.55
1:BBB:157:GLU:HB2	3:BBB:906:GOL:H2	1.88	0.55
1:AAA:102:ASP:OD2	4:AAA:403:NAD:H2N	2.07	0.54
1:BBB:7:ILE:H	1:BBB:256:ASN:ND2	2.04	0.54
1:BBB:45[B]:CYS:SG	1:BBB:47:VAL:HG23	2.49	0.53
1:AAA:218:HIS:CE1	1:AAA:253:GLY:HA3	2.43	0.53
1:AAA:34:ALA:HB2	1:AAA:91:TYR:CZ	2.46	0.51
4:AAA:403:NAD:C5N	5:AAA:719:HOH:O	2.59	0.50
1:BBB:34:ALA:HB2	1:BBB:91:TYR:CZ	2.46	0.50
1:AAA:31:TYR:HB3	1:AAA:91:TYR:CD1	2.47	0.49
1:AAA:346:ARG:NH1	5:AAA:501:HOH:O	2.15	0.48
1:AAA:157:GLU:OE1	3:AAA:405:GOL:H12	2.14	0.48
1:AAA:75:THR:HG1	3:AAA:405:GOL:C1	2.28	0.46
1:BBB:99:SER:HB2	1:BBB:100:PRO:HD3	1.98	0.46
1:AAA:99:SER:HB2	1:AAA:100:PRO:HD3	1.98	0.45
1:BBB:31:TYR:HB3	1:BBB:91:TYR:CD1	2.51	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:BBB:75:THR:HG1	3:BBB:906:GOL:C3	2.29	0.44
1:BBB:97:GLY:O	1:BBB:101[B]:GLN:HG2	2.18	0.44
1:BBB:218:HIS:HD2	5:BBB:1153:HOH:O	2.00	0.44
1:BBB:302:ASP:OD1	1:BBB:305:ARG:NH2	2.52	0.43
1:AAA:71:ASN:HD22	1:AAA:162:LYS:NZ	2.16	0.43
1:AAA:302:ASP:OD1	1:AAA:305[A]:ARG:NH2	2.50	0.43
1:AAA:218:HIS:NE2	1:AAA:253:GLY:HA3	2.33	0.43
1:BBB:291:TYR:HB2	1:BBB:378:LEU:HD21	2.00	0.43
1:AAA:251:GLY:HA2	1:AAA:254:ILE:HD12	2.01	0.42
1:AAA:45[B]:CYS:SG	1:AAA:47:VAL:HG23	2.60	0.41
1:AAA:151:ASN:HB2	4:AAA:403:NAD:N7N	2.35	0.41
1:AAA:157:GLU:HB2	3:AAA:405:GOL:H12	2.03	0.41
1:BBB:75:THR:OG1	3:BBB:906:GOL:H31	2.22	0.40
1:BBB:196:ASP:HB2	1:BBB:278:GLY:HA2	2.02	0.40
1:BBB:283:ILE:HA	3:BBB:908:GOL:H32	2.03	0.40
1:AAA:218:HIS:HD2	5:AAA:601:HOH:O	2.03	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	AAA	387/390 (99%)	381 (98%)	5 (1%)	1 (0%)	41 22
1	BBB	387/390 (99%)	382 (99%)	5 (1%)	0	100 100
All	All	774/780 (99%)	763 (99%)	10 (1%)	1 (0%)	51 31

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	254	ILE

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	AAA	298/299 (100%)	298 (100%)	0	100	100
1	BBB	298/299 (100%)	298 (100%)	0	100	100
All	All	596/598 (100%)	596 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

Mogul failed to run properly - this section is therefore empty.

5.5 Carbohydrates

Mogul failed to run properly - this section is therefore empty.

5.6 Ligand geometry

Mogul failed to run properly - this section is therefore empty.

5.7 Other polymers

Mogul failed to run properly - this section is therefore empty.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	AAA	383/390 (98%)	-0.13	10 (2%)	56	56	12, 18, 30, 42	0
1	BBB	384/390 (98%)	-0.31	6 (1%)	72	75	11, 17, 28, 50	0
All	All	767/780 (98%)	-0.22	16 (2%)	63	65	11, 18, 28, 50	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	385	SER	4.3
1	BBB	312	GLU	4.0
1	AAA	313	GLY	3.3
1	AAA	310	LYS	3.2
1	AAA	44	GLN	3.1
1	AAA	312	GLU	2.9
1	BBB	384	THR	2.8
1	BBB	313	GLY	2.8
1	AAA	385	SER	2.7
1	AAA	257	VAL	2.6
1	AAA	86	ASN	2.5
1	AAA	45[A]	CYS	2.5
1	BBB	386	HIS	2.5
1	AAA	237	LYS	2.4
1	BBB	341	ARG	2.4
1	AAA	37	VAL	2.3

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GOL	AAA	404	6/6	0.37	0.37	56,58,61,61	2
3	GOL	BBB	901	6/6	0.63	0.26	30,39,41,41	2
3	GOL	AAA	402	6/6	0.77	0.16	50,52,53,54	2
3	GOL	BBB	905	6/6	0.77	0.22	28,36,40,40	2
3	GOL	AAA	405	6/6	0.81	0.37	32,34,34,35	14
3	GOL	BBB	906	6/6	0.82	0.30	34,37,42,42	2
3	GOL	BBB	908	6/6	0.84	0.40	51,53,53,54	2
3	GOL	BBB	904	6/6	0.85	0.17	35,36,37,37	2
3	GOL	BBB	903	6/6	0.88	0.13	33,35,35,36	2
4	NAD	AAA	403	44/44	0.95	0.10	15,17,30,31	6
2	FE	AAA	401	1/1	0.98	0.08	18,18,18,18	0
4	NAD	BBB	907	44/44	0.98	0.09	13,16,23,25	6
2	FE	BBB	902	1/1	0.99	0.06	18,18,18,18	0

6.5 Other polymers (i)

There are no such residues in this entry.